Development and implementation of Novel Algorithms and Software Modules for PGR Informatics

Sunil Archak
ICAR–National Bureau of Plant Genetic Resources
Pusa Campus, New Delhi 110 012 INDIA

Disclaimer

This report contains the findings of the project "Development and implementation of Novel Algorithms and Software Modules for PGR Informatics". The findings represent the views of the National Fellow based on the analyses as contained in this report. No contents individually or in whole should be interpreted as the views of either ICAR or NBPGR.

IPRs

All the web apps and mobile apps described in this report are covered by the Copyright automatically after creation. However, the apps shall be registered as per the Copyright Act, 1957. ICAR National Fellow (with his team) is the Author of all the "software" developed. The ownership is assigned to ICAR, explicitly depicted in all apps. Data content is the property of ICAR-NBPGR.

Financial support

ICAR National Fellowship administered by the Agricultural Education Division of the Indian Council of Agricultural Research.

Project Details

Project No:	1006528
Project code:	16113130040
Scheme code:	90003
Duration:	60 months
Period:	Nov 2014 to Oct 2019
Budget released:	Rs. 170,00,000

Acknowledgements

The National Fellow gratefully acknowledges the constant guidance and strong support received from Prof. Kailash Bansal and Dr. Kuldeep Singh, past and present Directors of NBPGR. Contributions of NBPGR colleagues, present and past, to PGR documentation, are acknowledged. Significant contributions of Rajeev Gambhir, Vijay Mandal, Ratnesh K Tiwari and Firoz Ahmed are also acknowledged.

Contact information

Sunil Archak
ICAR National Fellow
NBPGR, Pusa Campus
New Delhi 110012
Tel: +91 11 2580 2883
Fax: +91 11 2584 2495
Email: sunil.archak@icar.gov.in

Correct citation

Sunil Archak (2019). Development and implementation of Novel Algorithms and Software Modules for PGR Informatics: Final Consolidated Report. National Bureau of Plant Genetic Resources, New Delhi. India. pp. 27.

Available online at:
http://pgrinformatics.nbpgr.ernet.in/download.aspx

Contents

1. Prelude

An organized digital information system provides fair and just opportunity for all to access Plant Genetic Resources (PGR). Enhanced access to information leads to enhanced PGR utilization. The ICAR-NF project on "Development and Implementation of Novel Algorithms and Software Modules for PGR Informatics" has aided the transformation of routine PGR documentation activity into PGR Informatics science. The Project has resulted not just in academic exercise of algorithms and prototypes but has generated a dozen unique and functional web and mobile apps for PGR Informatics and PGR Analytics. The novelty and utility of the web apps have seen them accessed by researchers across the globe. Impetus provided by the web-apps and mobile-apps developed in the NF Project to the use of PGR Portal by researchers in many countries demonstrates the immense impact of the ICAR-NF project. The Project has also facilitated in teaching and capacity building as well as in building infrastructure.

ICAR National Fellowship on PGR Informatics was a substantial and decisive step by ICAR towards scientific and institutionalized management of PGR information. This Progress Report entails brief description of the various informatics and analytics apps developed in the Project. More information on the PGR Informatics Project can be accessed at: http://pgrinformatics.nbpgr.ernet.in/

9th Aug 2019

Sunil Archak

NBPGR, New Delhi

2. The Project

Purpose of PGR Informatics

It is needless to emphasize the dependence of varietal development programs on the availability of trait specific plant genetic resources (PGR). However, PGR utilization depends upon the availability of information. The information on the available PGR must be compiled, curated and made accessible to users. Such information should be up to date, accurate, reliable, systematically stored and easily accessible. PGR information system must contain data on accessions (passport), augmentation (source and associated information), conservation (mode of conservation, viability, regeneration, etc.), traits (characterization and utilization), quarantine (phyto-sanitary status), utilization (supply and feedback). Management such diverse data types is a pre-requisite for (i) Genebank management; (ii) Utilization by researchers; and (iii) Compliance to international treaties/conventions and national legislations.

ICAR-NBPGR is the second largest genebank in the world conserving more than 400,000 accessions belonging to nearly all the cultivated species. However, lack of an operational online open access portal has led to (i) redundant investments in PGR characterization and evaluation experiments; (ii) inefficient utilization of valuable genetic resources; and (iii) a situation where policy makers and funding agencies remain uninformed about the exceptional progress made by Indian researchers. Primary reason for the absence of an effective information system was inadequate scientific, technical and financial support rendering the whole activity into mere routine *documentation* rather than an *informatics* system.

PGR Informatics, an emerging interdisciplinary science, is the application of information and communication technologies to collect, organize and analyze biological and environmental data about germplasm collections including phenotyping, genotyping, remote sensing, and phylo-climatic modelling. PGR Informatics manages (creation, storage, retrieval and presentation) and analyses (discovery, exploration and extraction) diverse information (facts, figures, statistics, knowledge and news) related to various aspects of augmentation, conservation, evaluation and utilization of plant genetic resources. Aim of PGR Informatics is to enhance the accessibility to relevant PGR information to enhance PGR utilization.

Need for research in PGR Informatics

PGR Informatics can only be built upon sound scientific grounds which go beyond computer expertise to encompass genetic, taxonomic, geo-informatic, bioinformatic and genomic linkages. A user-centric approach is needed that includes (i) Growth – through richer content and metadata especially field genebanks, horticultural genetic resources and farmers' efforts; (ii) Scalability – through established architecture and decentralised services; (iii) Capacity – through development of appropriate tools, products and services; (iv) Visibility – through development of a powerful and user-friendly data portal, nodes, thematic portals and rich internet applications.

Research and development in PGR analytics will accelerate cognitive decisions that facilitate PGR utilization. PGR analytics, at varying levels of analysis and complexity, includes (i) Descriptive - reporting and querying of data to identify problems and solutions; (ii) Predictive - modelling, forecasting, and simulating outcomes based on the data; and (iii) Prescriptive - recommend the best course of action based on the data.

Above-mentioned developments are not possible without laying down data standards as well as benchmarks for algorithms and modules. These in turn can't be developed without creating sufficient knowledge body based on basic research in PGR Informatics.

Globally, The genebank information systems are limited to 11 CGIAR genebanks and seven major national genebanks and is implemented at multiple levels: (i) Genebank level (e.g. CGN-PGR, Netherlands; IPK, Germany; CRF, Spain; VIR, Russia; EMBRAPA, Brazil; Kew's, UK; NIAS, Japan); (ii) National level (e.g. AusPGRIS, Australia; PGRC, Canada; NORDGEN; ARS-GRIN, USA); (iii) Regional level (e.g. EURISCO, Europe; EAPGREN, Eastern African Nations; REMERFI, Latin America; SPGRC, South African Countries; GRENEWECA, West and Central Africa); and (iv) Global level (System-wide Information Network for Genetic Resources, SINGER and GENESYS; Global Biodiversity Information Facility, GBIF). USDA genebank has developed an information management system called GRIN Global in 2008 (collaborative efforts of GCDT, Bioversity, USDA-ARS at $2.5 million). In 2010, GeneSys, a web portal for all the CG genebanks as well as EURISCO information system was launched. Unfortunately, at the other end of the spectrum, almost all the gene rich countries absolutely lack informatics facilities.

NBPGR is the nodal organization for PGR management in NARS. The ICAR National Fellow Project on "Development and implementation of Novel Algorithms and Software Modules for PGR Informatics" was implemented at NBPGR with effect from November 2014 in order to create nucleus of basic research in the field of PGR Informatics.

Critical gaps

A number of tools and algorithms have been developed in the field of informatics applicable in distinct areas like banking, clinical studies, engineering, social sciences and forensics. Some examples in biology include tools like Darwin Core Archive and communication protocols namely DiGIR, BioCASE, TAPIR–TDWG, etc. However, customizing available algorithms for PGR informatics continues to remain a challenge.

(i) Germplasm information management suffers not only because of inadequate digitization, collation and curation of data but also because the neither the information is standardized nor the compatible software and data models implemented

(ii) Absence of all-encompassing scientific questions to guide PGR informatics is resulting in developments that have no connection to genuine insight and forward progress. Researching on the links between diverse dimensions of plant genetic diversity– conservation status, taxonomy, distributional biology, ecology, interactions, genomics, and phylogenetics– will enable a transition from data management to informatics science

(iii) Lack of emphasis on PGR Analytics to facilitate cognitive decisions

(iv) Typically, data availability and technology have driven many of the ideas and concepts in the PGR informatics field. The desirable state of the field would see ideas and concepts driving development of new technology and new data resources.

Objectives of the Project

1. Development of data standards and algorithms for PGR informatics
2. Development of algorithms and modules for PGR analytics and bioinformatics
3. Development of algorithms and modules for user-centric approach to PGR informatics to encompass genetic, genomic, taxonomic, geo-informatic, bioinformatics, etc. linkages with genebank information

Technical Programme (with activity milestones and time-frame):

	Activity Milestone	Year1		Year2		Year3		Year4		Year5	
1.	Commissioning of the infrastructure and personnel										
2.	Algorithmic analysis, design, application development and testing										
3.	Development of software modules for PGR Informatics										
4.	Development of software modules for PGR Analytics										
5.	Development of a central PGR portal										
6.	Development of software modules for retrieving data from multiple databases										
7.	Development of bioinformatics modules for analysis related to PGR										
8.	Report writing, Publications of manuals and books										

Expected output(s) particularly on time scale (please quantify):

- Novel algorithms[1] for PGR Informatics and Analytics
- Software modules for PGR Informatics
- Software modules for PGR Analytics
- Central PGR portal
- Bioinformatics modules for analysis related to PGR

The following sections report the major output of the PGR Informatics project in terms of web-apps, mobile-apps and PGR Portal as well as the outcome and impact of the project.

[1] Please note that PGR algorithms are unlike statistical algorithms of machine learning, clustering, etc. and also unlike graph and greedy algorithms of computational biology. PGR informatics algorithms are non-canonical problem-solving computer operations with rules set by biology.

3. PGR Informatics Applications

A genebank carries out numerous disparate activities in no particular sequence owing to gap between successive operations. All the activities revolve around conservation. Suitable informatics intervention has the potential to draw intelligent connotations by linking the data generated by each of the activities *a priori* and *a posteriori* (Figure 1). PGR Informatics applications were designed and developed based on this principle.

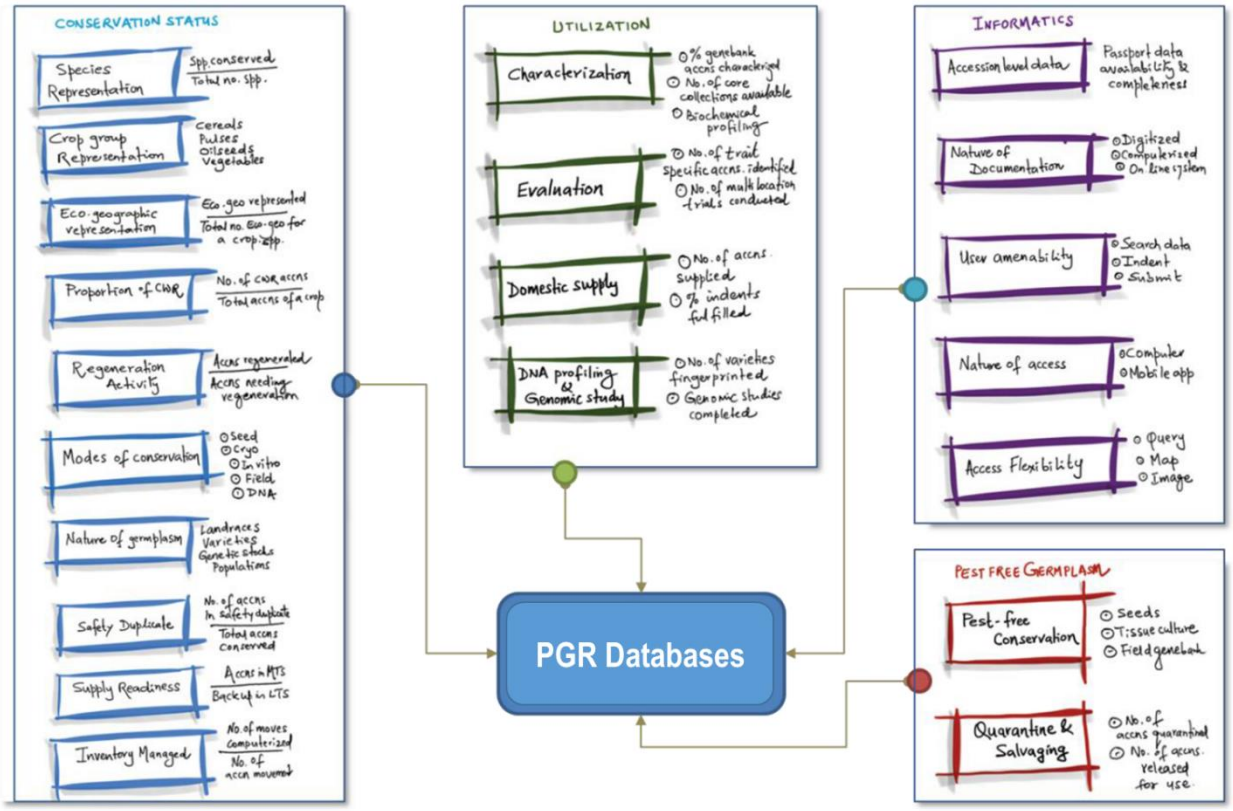


Figure 1. Elements and complexities of PGR management activities

The generic design of all PGR Informatics applications:

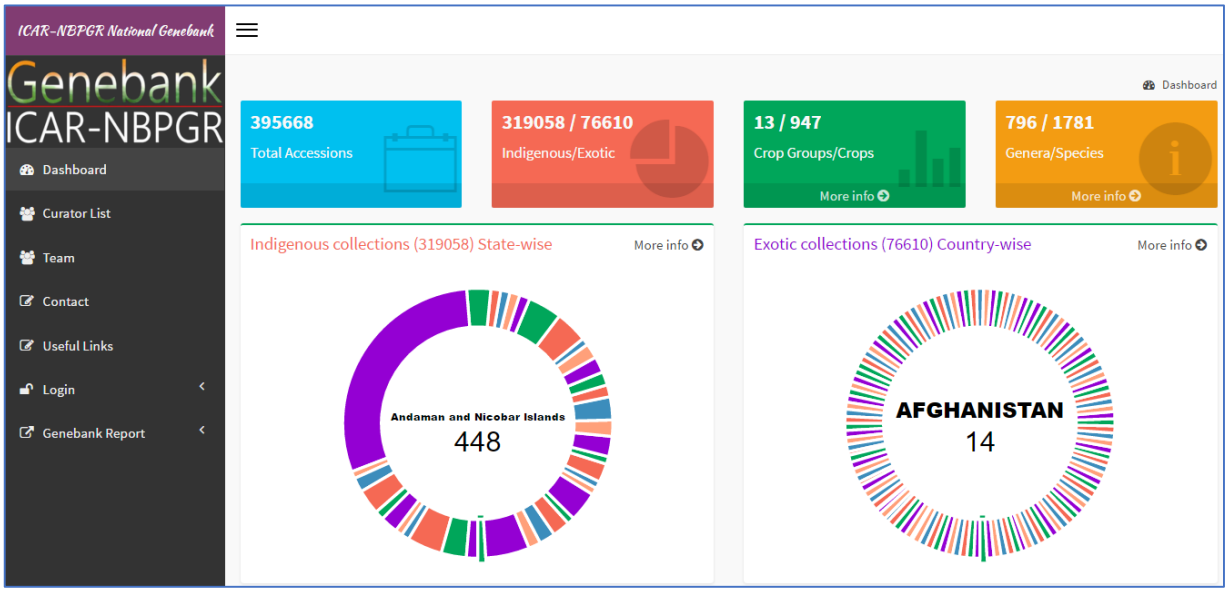
1. The databases and applications were built in English
2. Databases were built on MS SQL and applications were developed using ASP.NET with additional features wherever necessary
3. Map based depictions utilized Google map features (ensuring Indian boundaries are properly displayed)
4. Infographics provide salient information graphically; 2-3 clicks take users to detailed data (as much as intended in each application); structured reports (conceived based on stakeholder's demand) generated.
5. Performance of the applications was independent of scale, browser, location

Applications related to Conservation

Conservation activities are completely restricted to NBPGR and her regional stations. Germplasm is conserved either in seed genebank or as tissue cultured propagules in in vitro genebank or in the cryobank. Infrastructure, expertise, personnel, procedures, data required and generated, etc. are different in these conservation methods. Therefore, independent applications were developed for each of them.

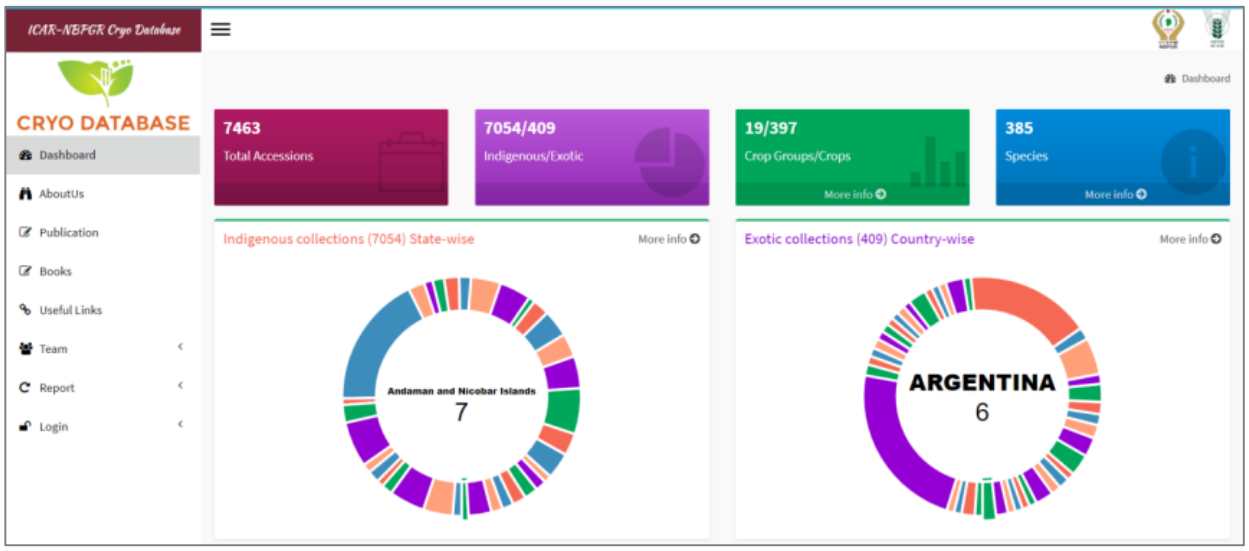
Seed Genebank (**genebank.nbpgr.ernet.in**)

A quick access to the information on number of germplasm accessions conserved in the National Genebank has been a long pending demand. A user friendly Genebank Dashboard was designed, developed and implemented to provide a personalized quick access to genebank information to PGR workers, breeders, students, research managers and administrators. The dashboard is designed to be compatible with old and new computers, tablets and smart phones. The dashboard figures are dynamically updated as and when genebank database is updated.



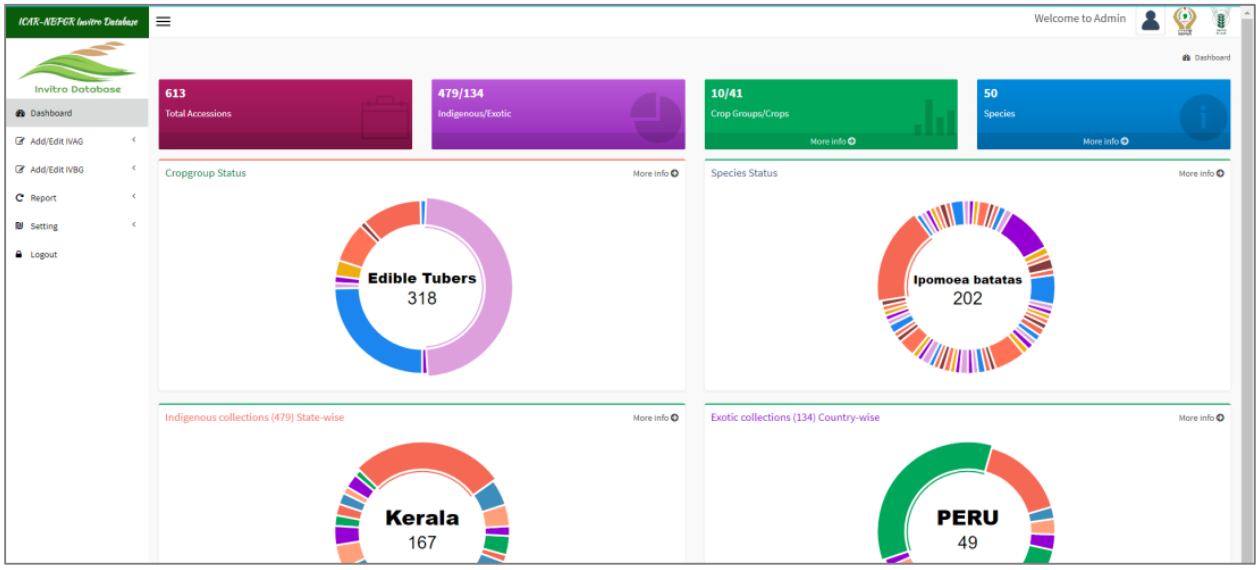
Cryo Database (pgrinformatics.nbpgr.ernet.in/cryobank)

Cryobanking started at NBPGR in 1960. Currently NBPGR cryobank is the 3rd largest in the world conserving 11,680 accessions of 794 species. This is a multi-crop repository unlike many crop-specific cryobanks in the world. Now with availability of an online dashboard everyone can find the status of the conservation with the help of infographics. The application also has made available related protocols developed and published as peer reviewed papers and books. It is expected that researchers around the world will benefit from the information and much needed impetus to cryobanking will come forth.



In vitro Database (pgrinformatics.nbpgr.ernet.in/invitro)

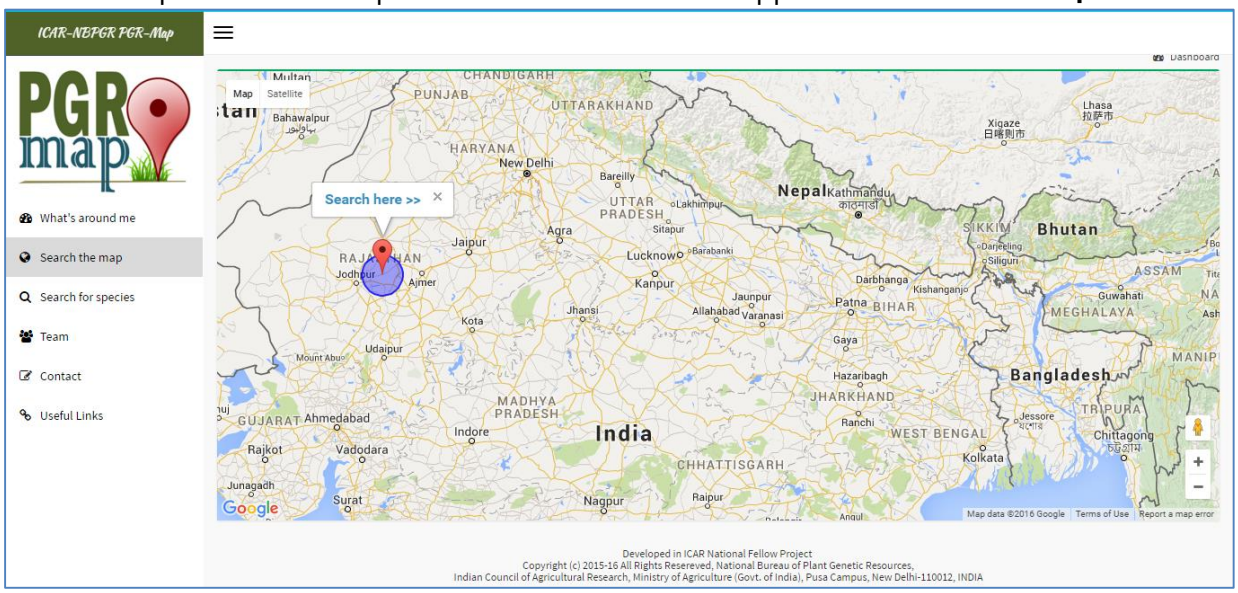
A useful complement to the conventional seed genebanking, in vitro conservation has an important role in using biotechnological approaches for conservation of PGR. A total of 40,000 cultures of 1890 accessions belonging to 133 species of fruits, tuber and bulbous crops, spices, plantation and industrial crops, medicinal, aromatic and rare/endangered plants are maintained in vitro, under culture room conditions or at low temperature.



Applications related to Exploration

PGR Map (http://pgrinformatics.nbpgr.ernet.in/pgrmap/)

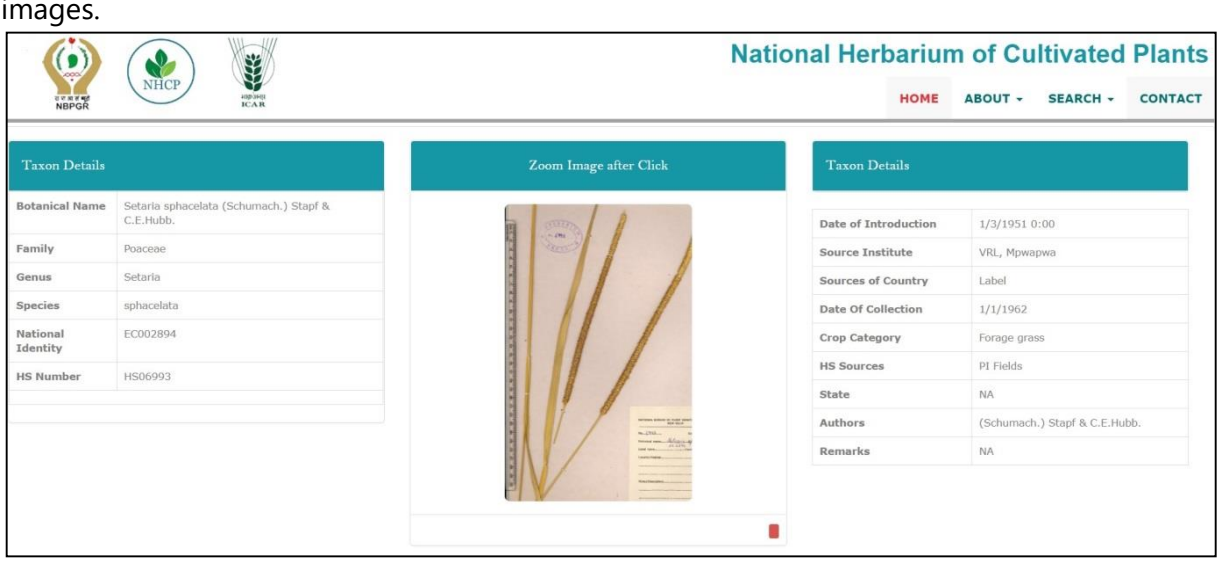
Map-based data retrieval provides easy and intuitive access to PGR information. The process is vital for developing mobile apps. Computational algorithms for map based applications in PGR were developed and were implemented as an interactive application called "**PGR Map**".



PGR Map offers three benefits: (i) "What's around me" helps user to obtain quickly the accessions that have been collected and conserved in the genebank from a particular location in India where the user is located at the moment; (ii) "Search the map" helps user to list the accessions that have been collected and conserved in the genebank from any selected location in India; (iii) "Search for species" helps user to map the collection sites of a crop species.

Virtual Herbarium of Crop Plants (http://pgrinformatics.nbpgr.ernet.in/nhcp)

National Herbarium of Crop Plants (NHCP), established in 1985 at NBPGR, has 23,665 specimens representing 267 families, 1,521 genera and 4,271 species. In order to make students and researchers across the world to access the invaluable information by the click of a button, ICAR-NF collaborated with NHCP to develop an online application creating a virtual herbarium with over 3,500 species of crop genepools complete with taxonomic information and about 7,000 images.



Applications for Intellectual Property related to PGR

Germplasm Registration Information System (www.nbpgr.ernet.in/registration)

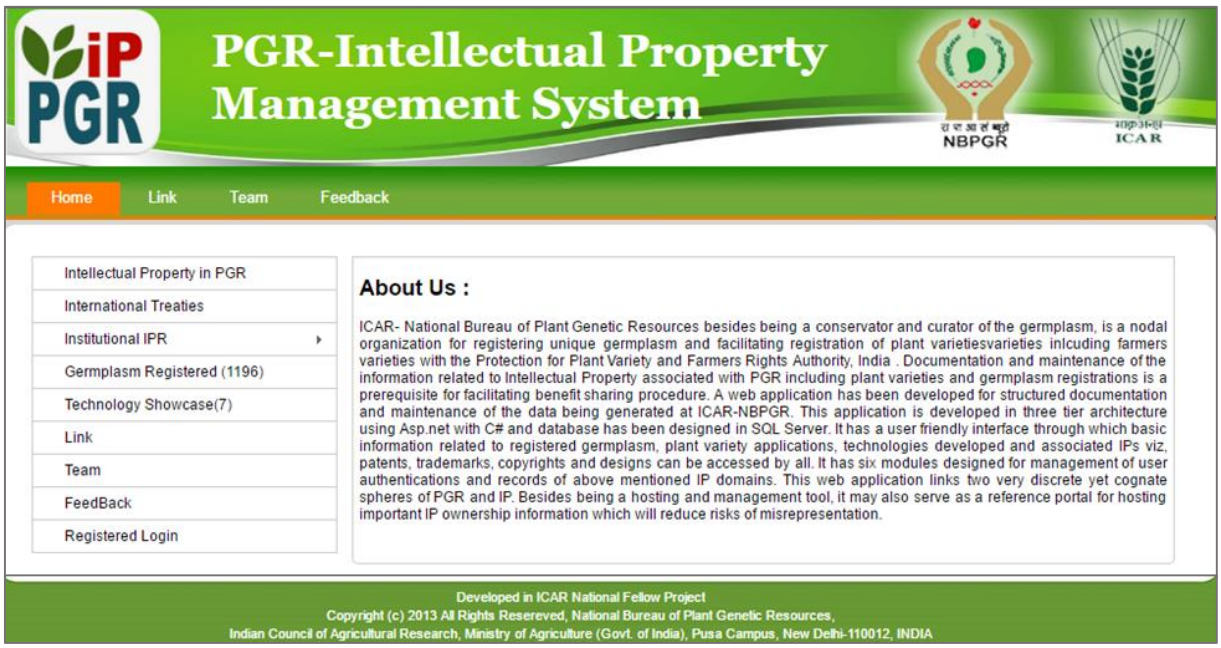
Plant Genetic Resources for Food and Agriculture form the basis for development of new varieties. In order (i) to recognize the developers of unique potentially valuable germplasm and genetic stocks and (ii) to inventorize, document and bring all the important genetic resources into public domain, facilitating their safe and accelerated use in research and crop improvement, a mechanism for "Registration of Plant Germplasm" was instituted in 1996 at NBPGR, New Delhi by the ICAR.

The Germplasm Registration Information System has been developed to make the entire process of germplasm registration—submission of application, evaluation by experts and decision by Plant Germplasm Registration Committee—a web-based system. The system is expected to provide genebank managers, breeders and plant researchers with a hands-on tool for management of germplasm registration process, and to policy makers with a reliable source of information. With the advent of this system, it is expected that the entire process of germplasm registration is made simple, transparent and fast.

Since its launch, germplasm registration activity is going exclusively through the GRIS and certificates can be downloaded by the scientists any time.

PGR IP-Management System (pgrinformatics.nbpgr.ernet.in/ip-pgr)

Systematic documentation and management of information related to IPs associated with PGR is a prerequisite for facilitating benefit sharing mechanism. This application facilitates access to various stakeholders.



Applications for Germplasm Exchange and Supply

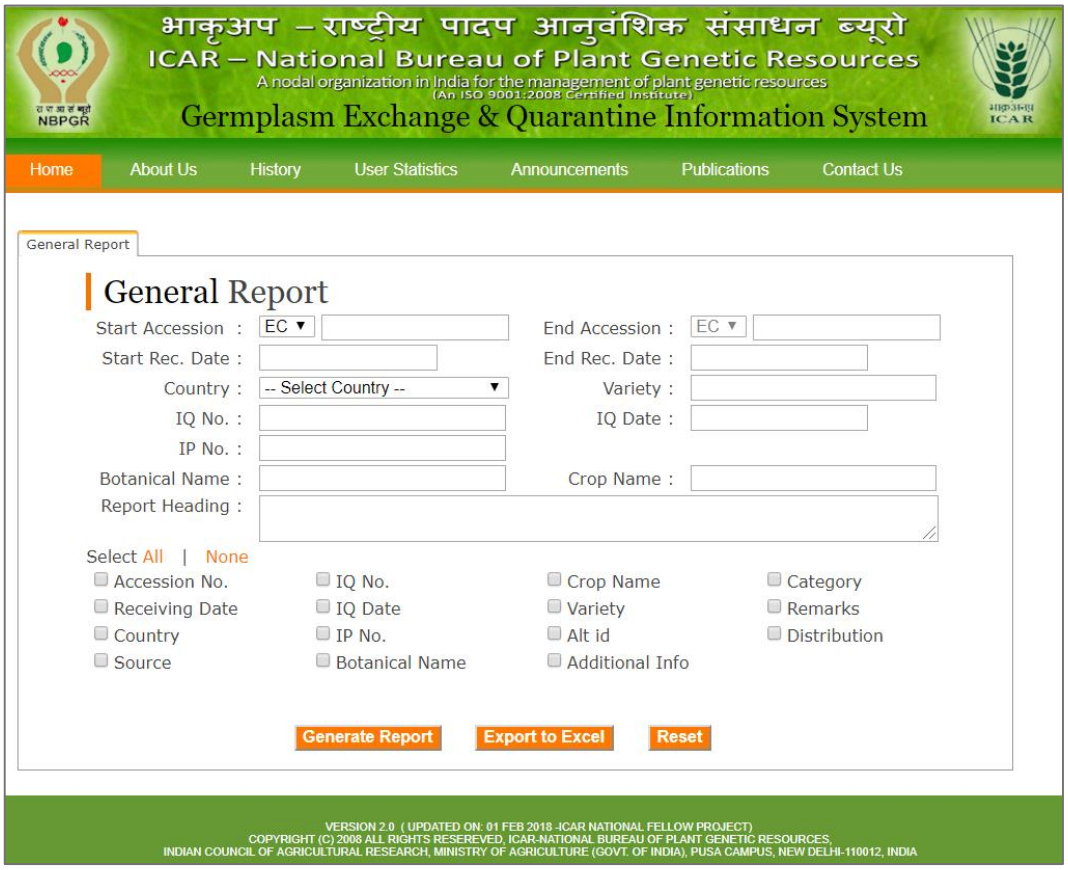
NBPGR is the nodal organization for import of PGR for research purposes. Researchers and companies importing germplasm register themselves using the online system and place indent for germplasm. The process is handled using the web application.

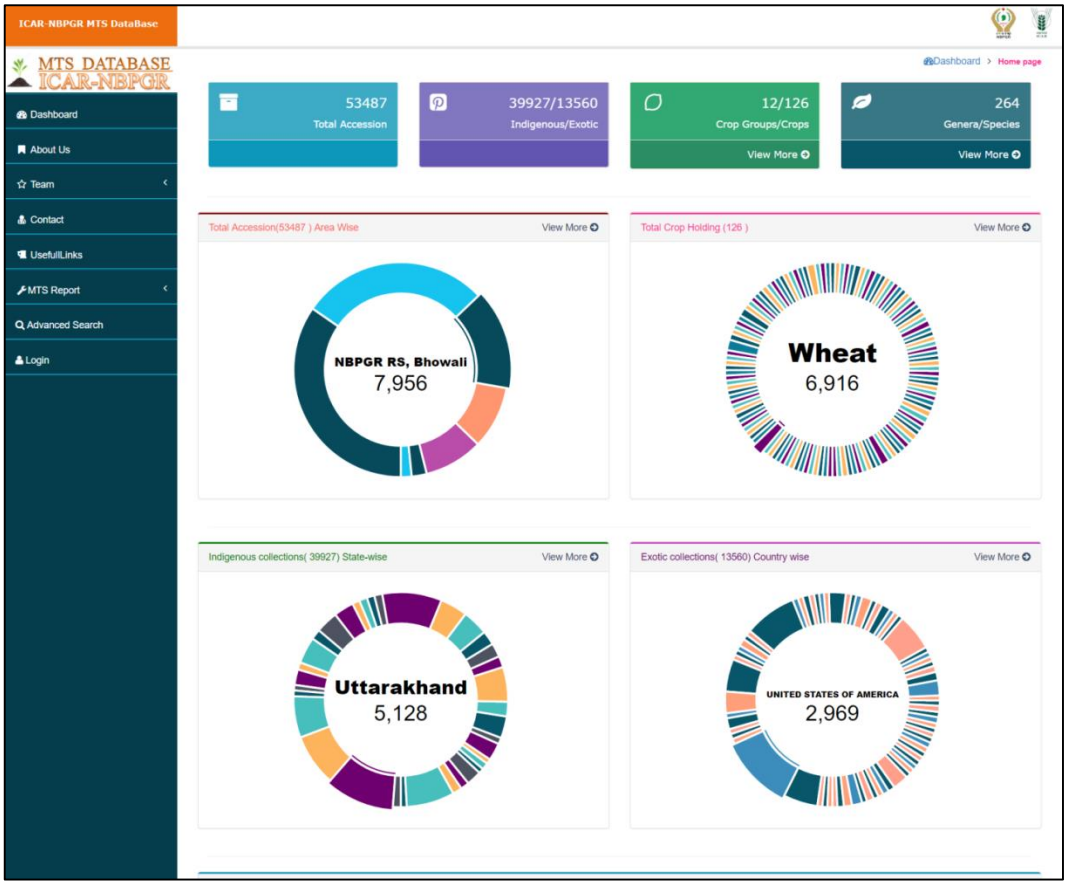
Germplasm Exchange (exchange.nbpgr.ernet.in)

Germplasm Exchange Unit caters to the germplasm needs of different researchers/scientists working in various research organizations in the country. Germplasm exchange activities at ICAR-NBPGR have major responsibility of Introducing (Importing) PGR; Facilitating export of PGR and arranging for National supplies of PGR.

Germplasm Supply (pgrinformatics.nbpgr.ernet.in/mts)

Utilization of PGR is the central focus of PGR management activities. Germplasm accessions are conserved in medium term storage modules (unlike genebank) that exist at NBPGR and many of her regional stations as well as some crop based IACR institutes. Domestic supply of PGR is serviced through these centres with MTS facility to store active germplasm for immediate supply. Effective utilization starts with inventorization and a pipeline for online management of indents. An application dedicated to the supply was developed and under beta testing.

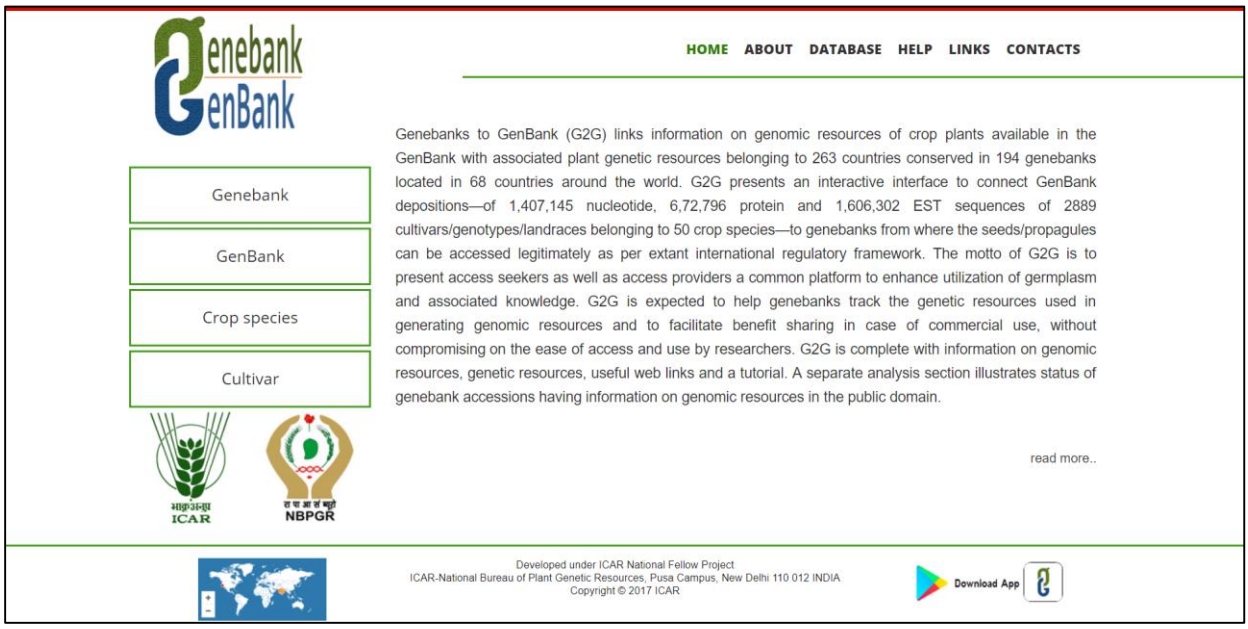

भाकृअप – राष्ट्रीय पादप आनुवंशिक संसाधन ब्यूरो
ICAR – National Bureau of Plant Genetic Resources
A nodal organization in India for the management of plant genetic resources
(An ISO 9001:2008 Certified Institute)
Germplasm Exchange & Quarantine Information System
Home
About Us
History
User Statistics
Announcements
Publications
Contact Us
General Report
General Report
Start Accession :
End Accession :
Start Rec. Date :
End Rec. Date :
Country :
-- Select Country --
Variety :
IQ No. :
IQ Date :
IP No. :
Botanical Name :
Crop Name :
Report Heading :
Select All | None
Accession No.
IQ No.
Crop Name
Category
Receiving Date
IQ Date
Variety
Remarks
Country
IP No.
Alt id
Distribution
Source
Botanical Name
Additional Info
Generate Report
Export to Excel
Reset
VERSION 2.0 (UPDATED ON: 01 FEB 2018 -ICAR NATIONAL FELLOW PROJECT)
COPYRIGHT (C) 2008 ALL RIGHTS RESEREVED, ICAR-NATIONAL BUREAU OF PLANT GENETIC RESOURCES,
INDIAN COUNCIL OF AGRICULTURAL RESEARCH, MINISTRY OF AGRICULTURE (GOVT. OF INDIA), PUSA CAMPUS, NEW DELHI-110012, INDIA

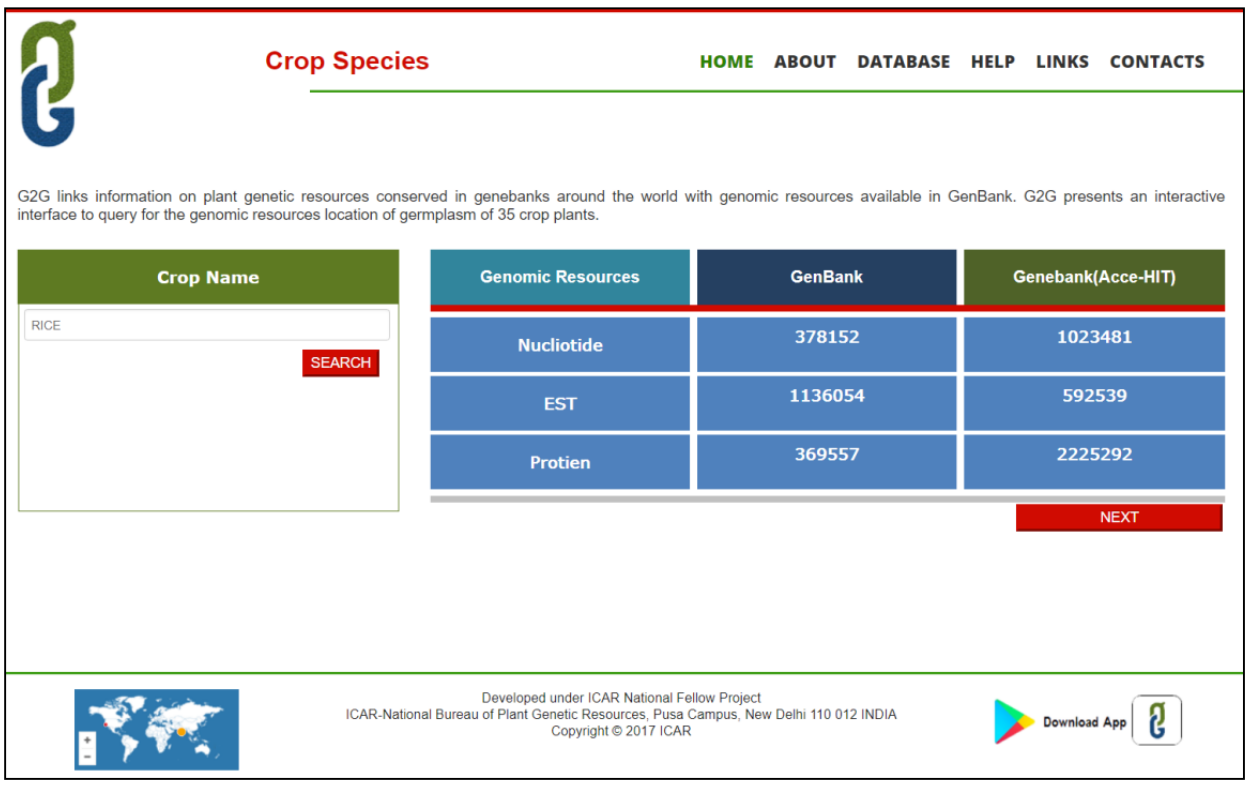

ICAR-NBPGR MTS DataBase
MTS DATABASE
ICAR-NBPGR
Dashboard
About Us
Team
Contact
UsefullLinks
MTS Report
Advanced Search
Login
Dashboard > Home page
53487
Total Accession
39927/13560
Indigenous/Exotic
12/126
Crop Groups/Crops
View More
264
Genera/Species
View More
Total Accession(53487) Area Wise
View More
NBPGR RS, Bhowali
7,956
Total Crop Holding (126)
View More
Wheat
6,916
Indigenous collections(39927) State-wise
View More
Uttarakhand
5,128
Exotic collections(13560) Country wise
View More
UNITED STATES OF AMERICA
2,969

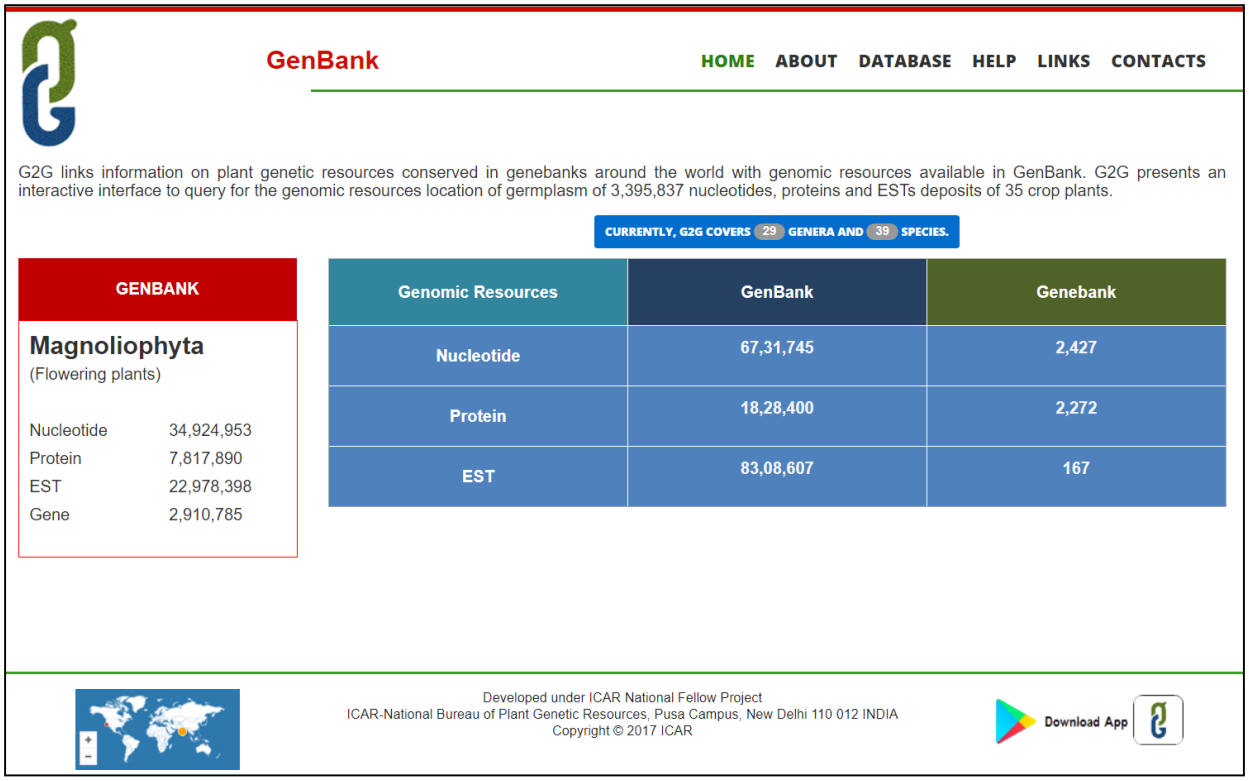
4. PGR Analytics Applications

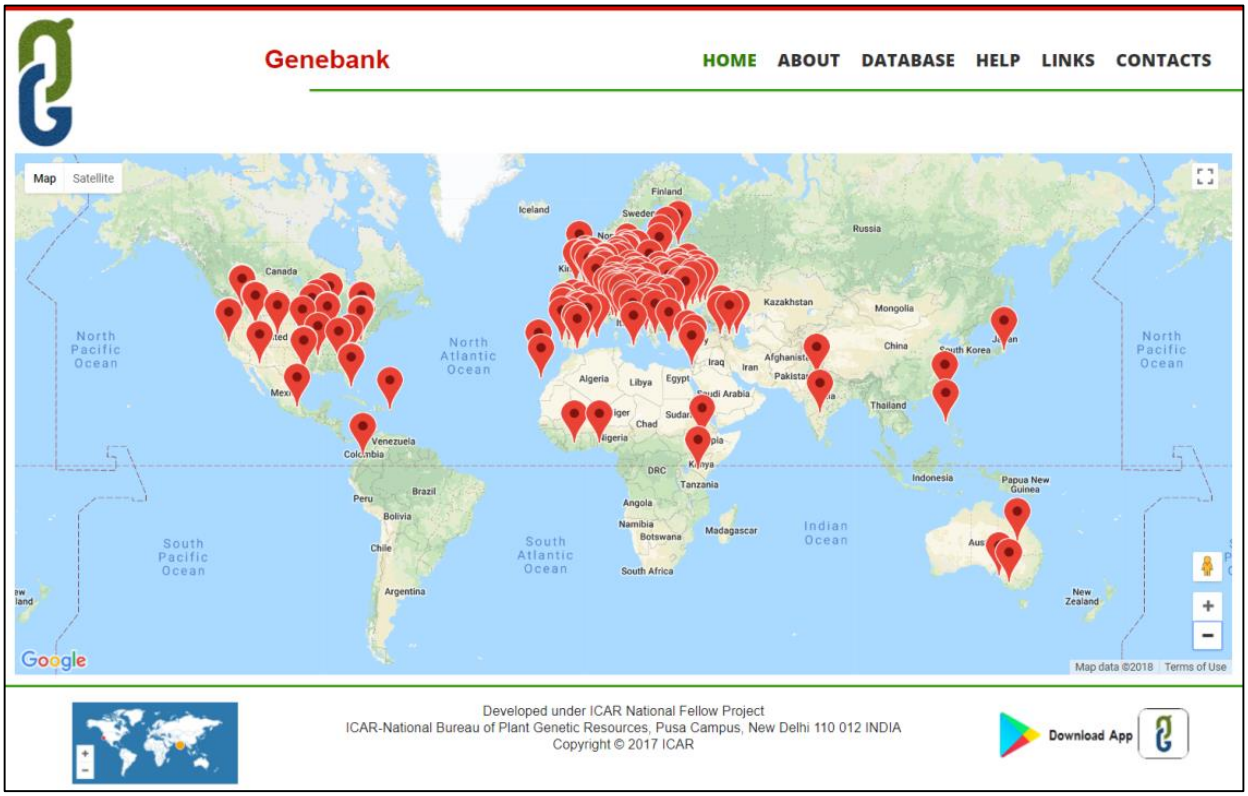
Genebank to Genbank G2G (http://pgrinformatics.nbpgr.ernet.in/g2g)

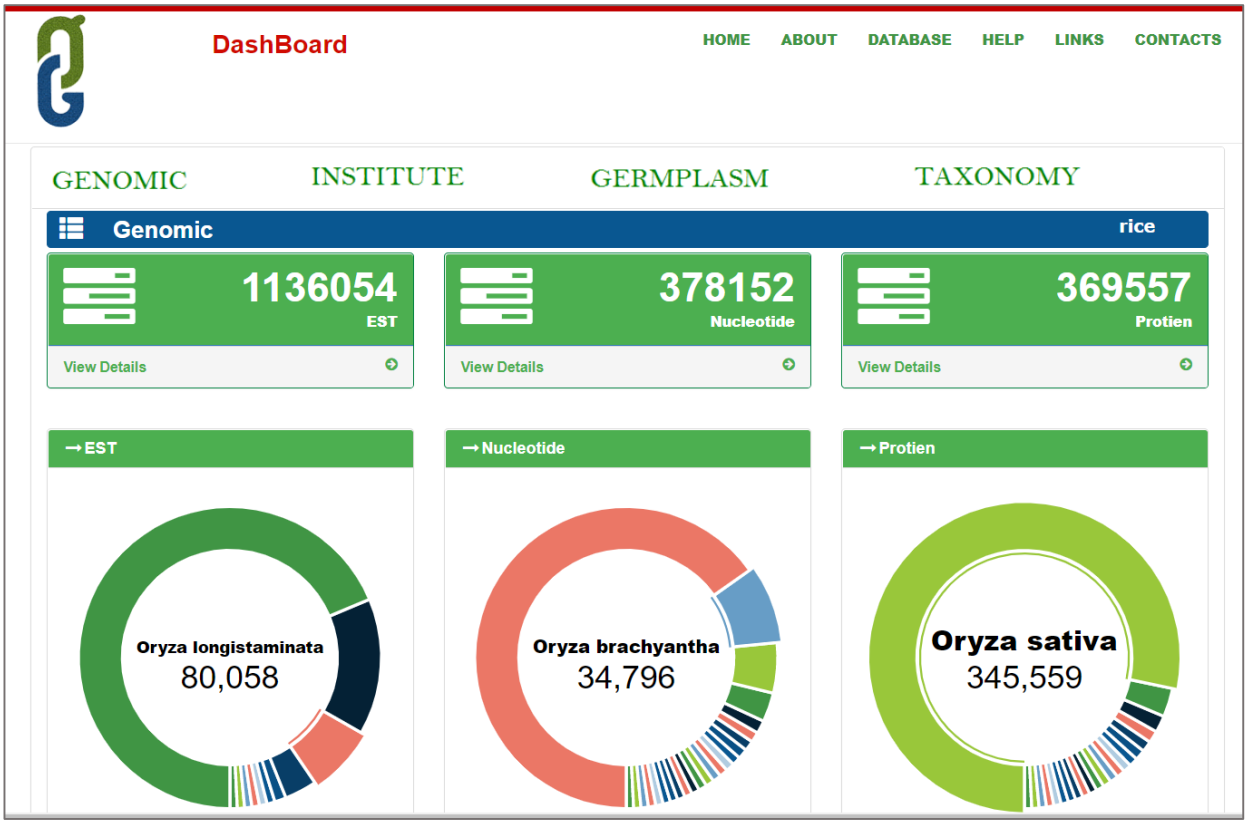
G2G links information on genomic resources of crop plants available in the GenBank with associated plant genetic resources belonging to 263 countries conserved in 194 genebanks located in 68 countries around the world. G2G connects GenBank depositions—of 1,407,145 nucleotide, 6,72,796 protein and 1,606,302 EST sequences of 2889 cultivars/ genotypes/ landraces belonging to 50 crop species—to genebanks from where the seeds/propagules can be accessed legitimately as per extant international regulatory framework. The motto of G2G is to present access seekers as well as access providers a common platform to enhance utilization of germplasm and associated knowledge.

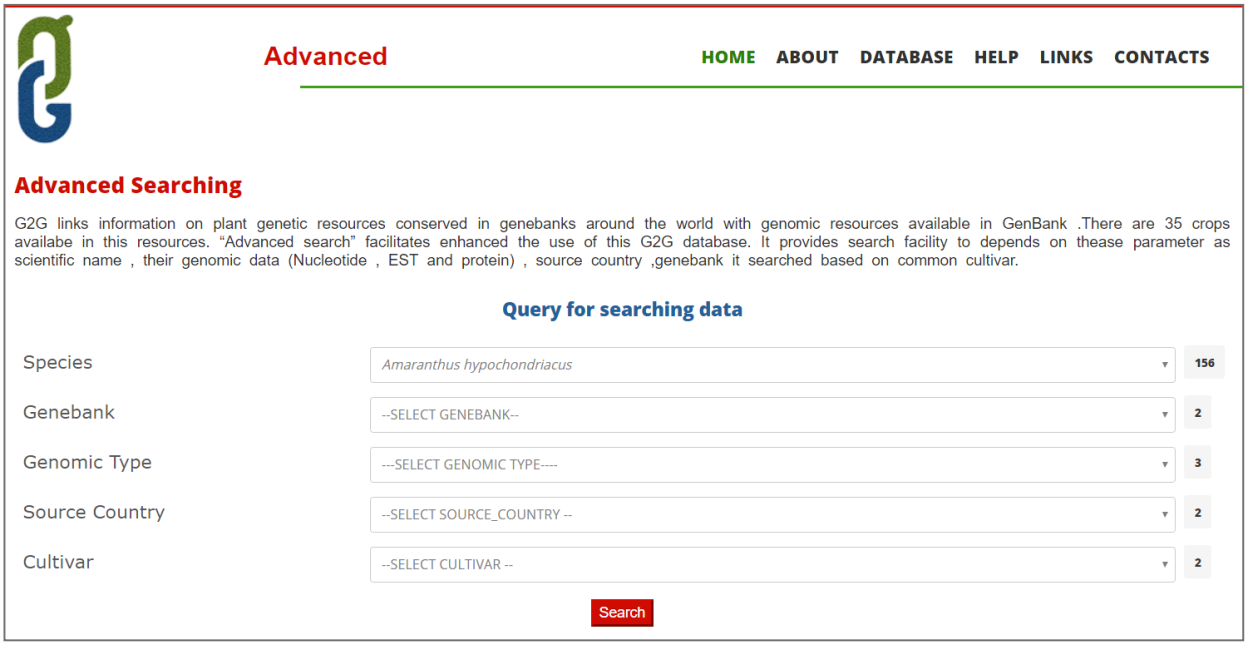


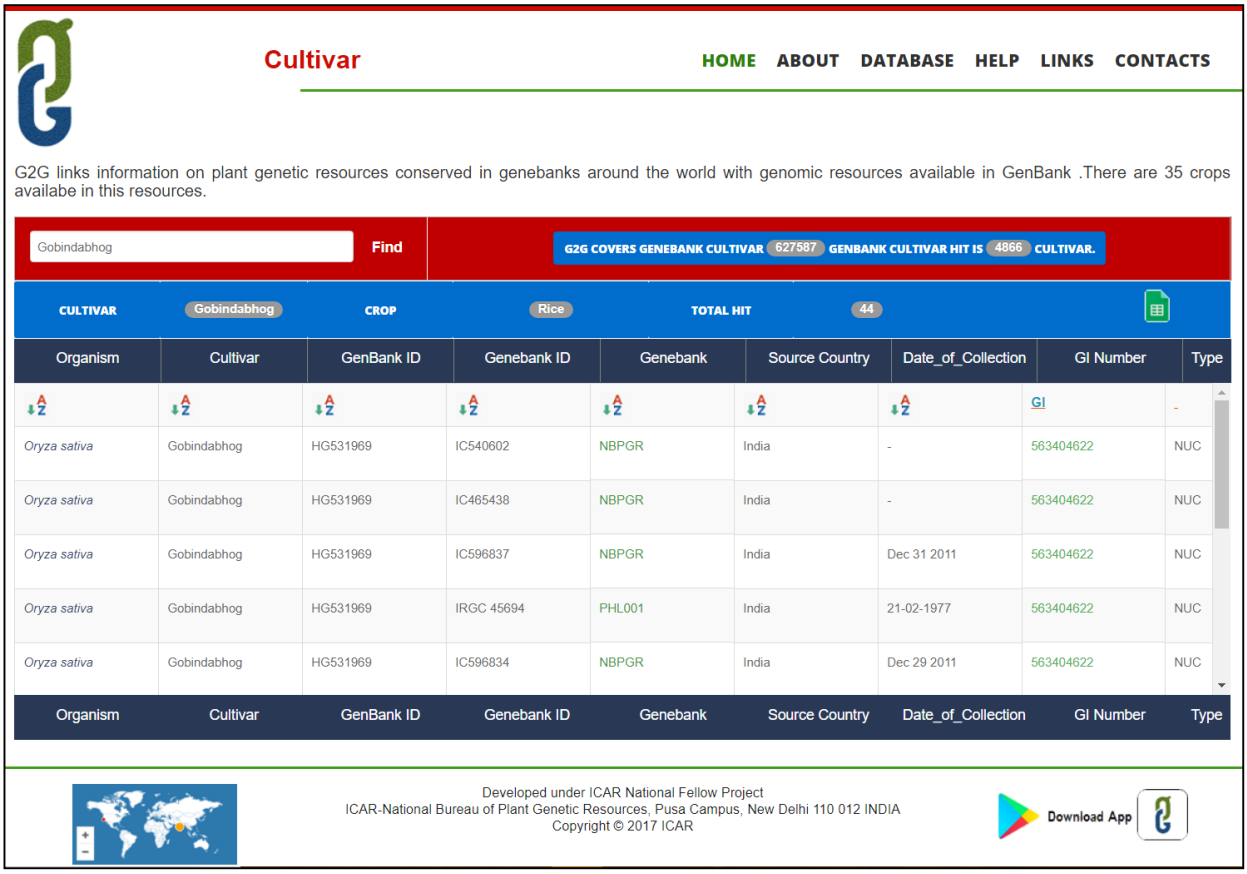


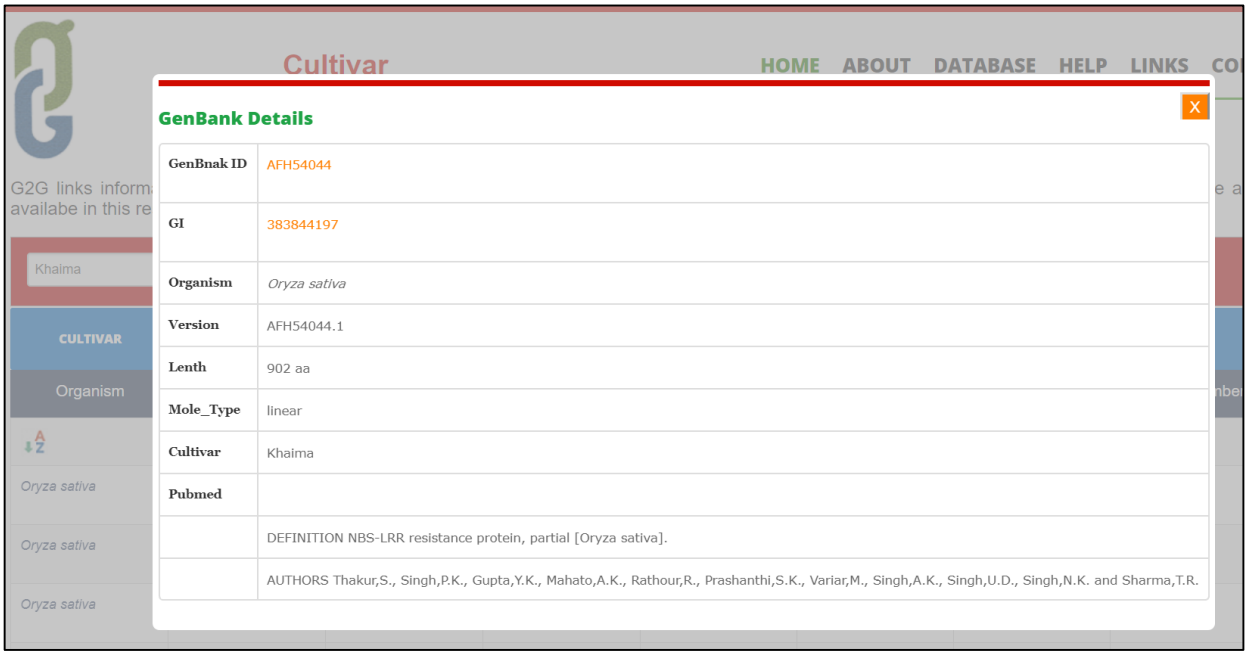

GenBank
HOME ABOUT DATABASE HELP LINKS CONTACTS
G2G links information on plant genetic resources conserved in genebanks around the world with genomic resources available in GenBank. G2G presents an interactive interface to query for the genomic resources location of germplasm of 3,395,837 nucleotides, proteins and ESTs deposits of 35 crop plants.
CURRENTLY, G2G COVERS 29 GENERA AND 39 SPECIES.
GENBANK
Magnoliophyta
(Flowering plants)
Nucleotide 34,924,953
Protein 7,817,890
EST 22,978,398
Gene 2,910,785
Genomic Resources | GenBank | Genebank
Nucleotide | 67,31,745 | 2,427
Protein | 18,28,400 | 2,272
EST | 83,08,607 | 167
Developed under ICAR National Fellow Project
ICAR-National Bureau of Plant Genetic Resources, Pusa Campus, New Delhi 110 012 INDIA
Copyright © 2017 ICAR
Download App


Genebank
HOME ABOUT DATABASE HELP LINKS CONTACTS
Map Satellite
North Pacific Ocean
North Atlantic Ocean
South Pacific Ocean
South Atlantic Ocean
Indian Ocean
Canada
Russia
Finland
Iceland
Kazakhstan
Mongolia
China
Thailand
Indonesia
Papua New Guinea
Australia
New Zealand
Algeria
Libya
Egypt
Chad
Nigeria
Sudan
Kenya
Tanzania
DRC
Angola
Namibia
Botswana
Madagascar
South Africa
Venezuela
Colombia
Peru
Brazil
Bolivia
Chile
Argentina
Iraq
Iran
Afghanistan
Pakistan
Saudi Arabia
Google
Map data ©2018
Terms of Use
Developed under ICAR National Fellow Project
ICAR-National Bureau of Plant Genetic Resources, Pusa Campus, New Delhi 110 012 INDIA
Copyright © 2017 ICAR
Download App


DashBoard
HOME ABOUT DATABASE HELP LINKS CONTACTS
GENOMIC INSTITUTE GERMPLASM TAXONOMY
Genomic
rice
1136054
EST
View Details
378152
Nucleotide
View Details
369557
Protien
View Details
EST
Oryza longistaminata
80,058
Nucleotide
Oryza brachyantha
34,796
Protien
Oryza sativa
345,559

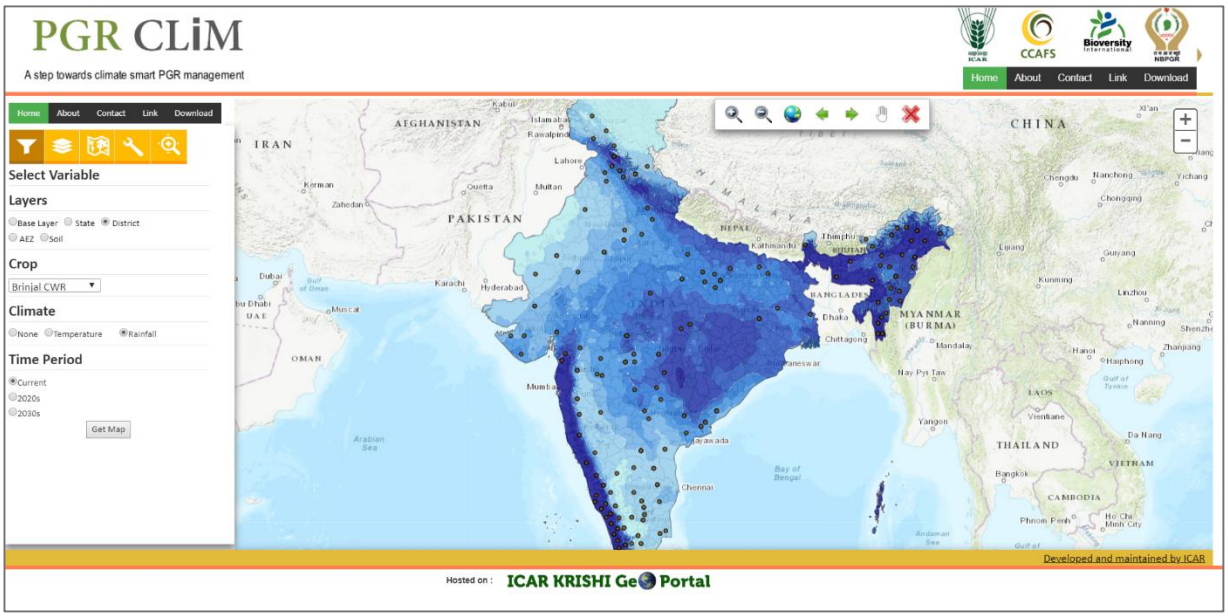

Advanced
HOME ABOUT DATABASE HELP LINKS CONTACTS
Advanced Searching
G2G links information on plant genetic resources conserved in genebanks around the world with genomic resources available in GenBank .There are 35 crops availabe in this resources. "Advanced search" facilitates enhanced the use of this G2G database. It provides search facility to depends on thease parameter as scientific name , their genomic data (Nucleotide , EST and protein) , source country ,genebank it searched based on common cultivar.
Query for searching data
Species
Amaranthus hypochondriacus
156
Genebank
--SELECT GENEBANK--
2
Genomic Type
---SELECT GENOMIC TYPE----
3
Source Country
--SELECT SOURCE_COUNTRY --
2
Cultivar
--SELECT CULTIVAR --
2
Search


Cultivar
HOME ABOUT DATABASE HELP LINKS CONTACTS
G2G links information on plant genetic resources conserved in genebanks around the world with genomic resources available in GenBank .There are 35 crops availabe in this resources.
Gobindabhog
Find
G2G COVERS GENEBANK CULTIVAR 627587 GENBANK CULTIVAR HIT IS 4866 CULTIVAR.
CULTIVAR Gobindabhog CROP Rice TOTAL HIT 44
Organism Cultivar GenBank ID Genebank ID Genebank Source Country Date_of_Collection GI Number Type
Oryza sativa Gobindabhog HG531969 IC540602 NBPGR India - 563404622 NUC
Oryza sativa Gobindabhog HG531969 IC465438 NBPGR India - 563404622 NUC
Oryza sativa Gobindabhog HG531969 IC596837 NBPGR India Dec 31 2011 563404622 NUC
Oryza sativa Gobindabhog HG531969 IRGC 45694 PHL001 India 21-02-1977 563404622 NUC
Oryza sativa Gobindabhog HG531969 IC596834 NBPGR India Dec 29 2011 563404622 NUC
Developed under ICAR National Fellow Project
ICAR-National Bureau of Plant Genetic Resources, Pusa Campus, New Delhi 110 012 INDIA
Copyright © 2017 ICAR
Download App


Cultivar
HOME ABOUT DATABASE HELP LINKS
GenBank Details
GenBnak ID AFH54044
GI 383844197
Organism Oryza sativa
Version AFH54044.1
Lenth 902 aa
Mole_Type linear
Cultivar Khaima
Pubmed
DEFINITION NBS-LRR resistance protein, partial [Oryza sativa].
AUTHORS Thakur,S., Singh,P.K., Gupta,Y.K., Mahato,A.K., Rathour,R., Prashanthi,S.K., Variar,M., Singh,A.K., Singh,U.D., Singh,N.K. and Sharma,T.R.

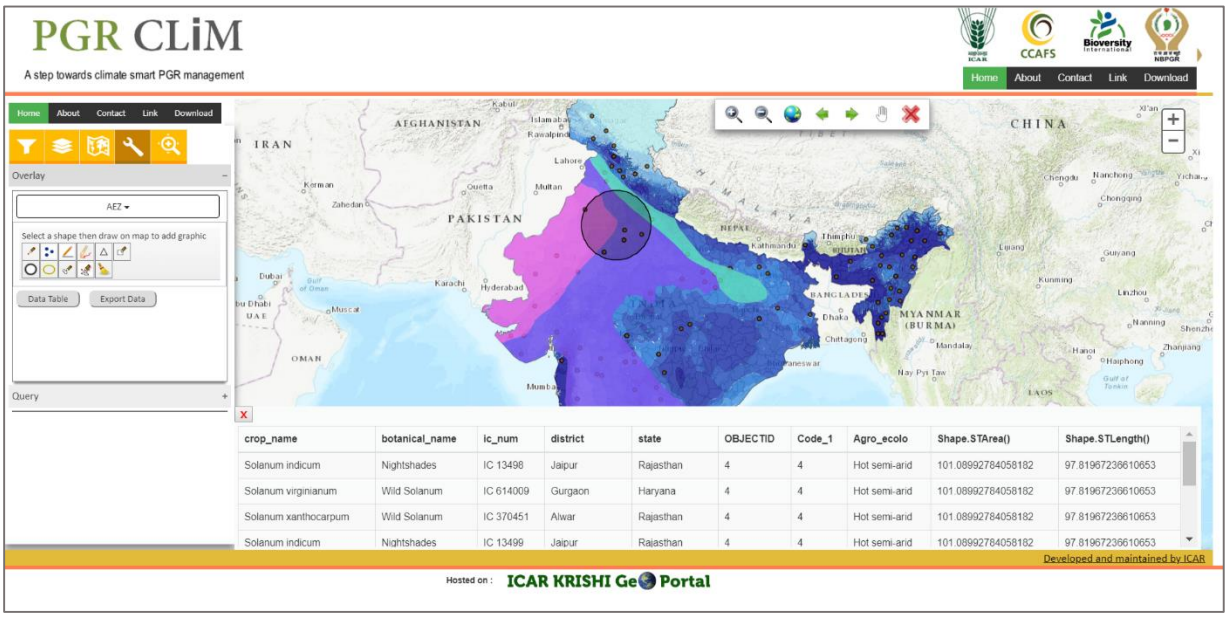
Geo-informatics portal in PGR (http://pgrinformatics.nbpgr.ernet.in/pgrclim)

A GIS-server based interactive application was implemented on ICAR Krishi Geo Portal comprising of layers of germplasm accessions of ten major crops; soil type; AEZ; temperature and rainfall (current, 2020 and 2030). The data for the PGR Clim was generated in a CCAFS funded project to link germplasm to changing climatic regimes.



Users can query to geographically locate about 10,000 accessions belonging to ten crops along with political boundaries (state/district), AEZ, soil type, climatic parameters (temperature or rainfall) for current and predicted future. High quality maps can be generated to be used by researchers for their analyses.

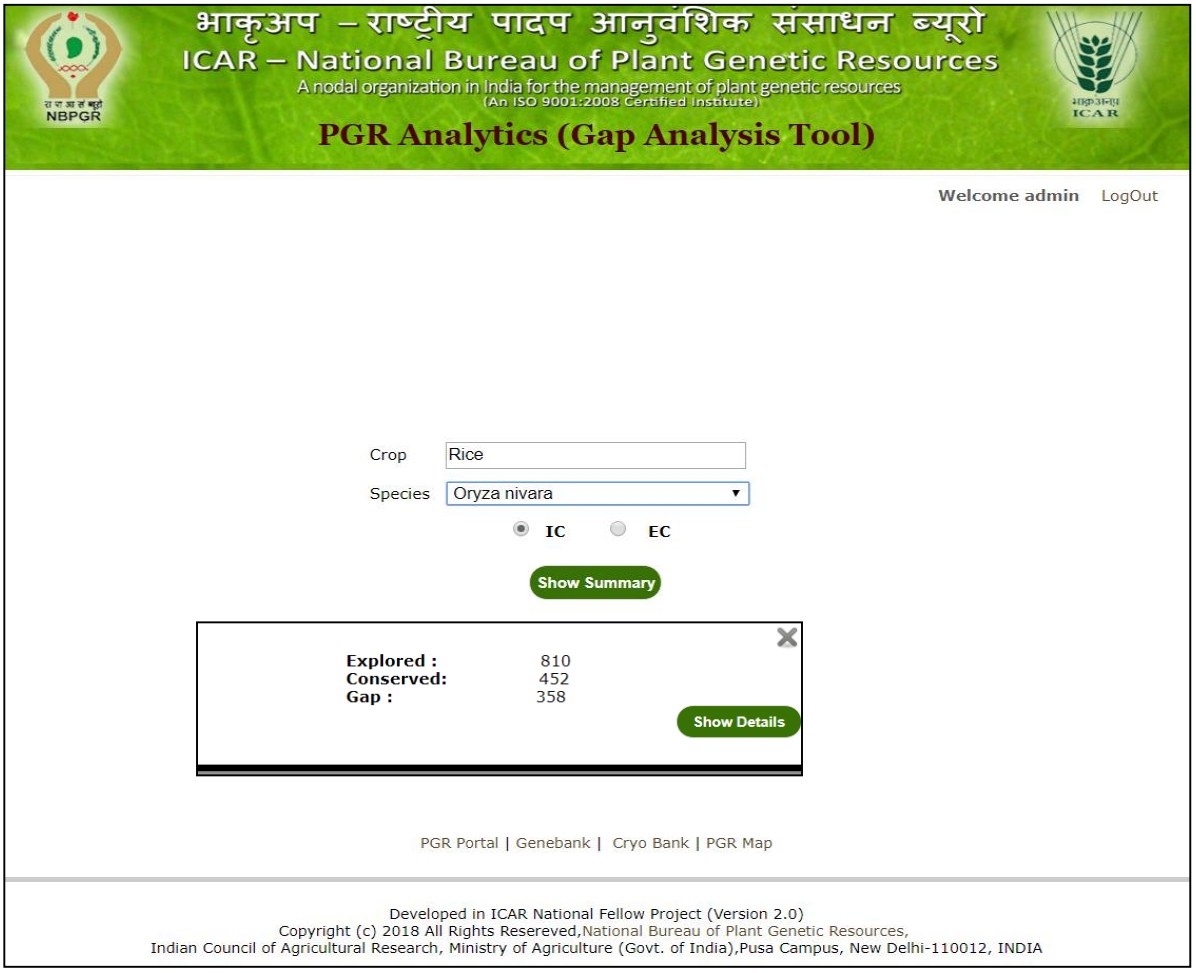
Additionally, results can be exported as excel files for further use by the researchers.



Gap Analysis Version 2.0 (Meant for internal use of NBPGR scientists)

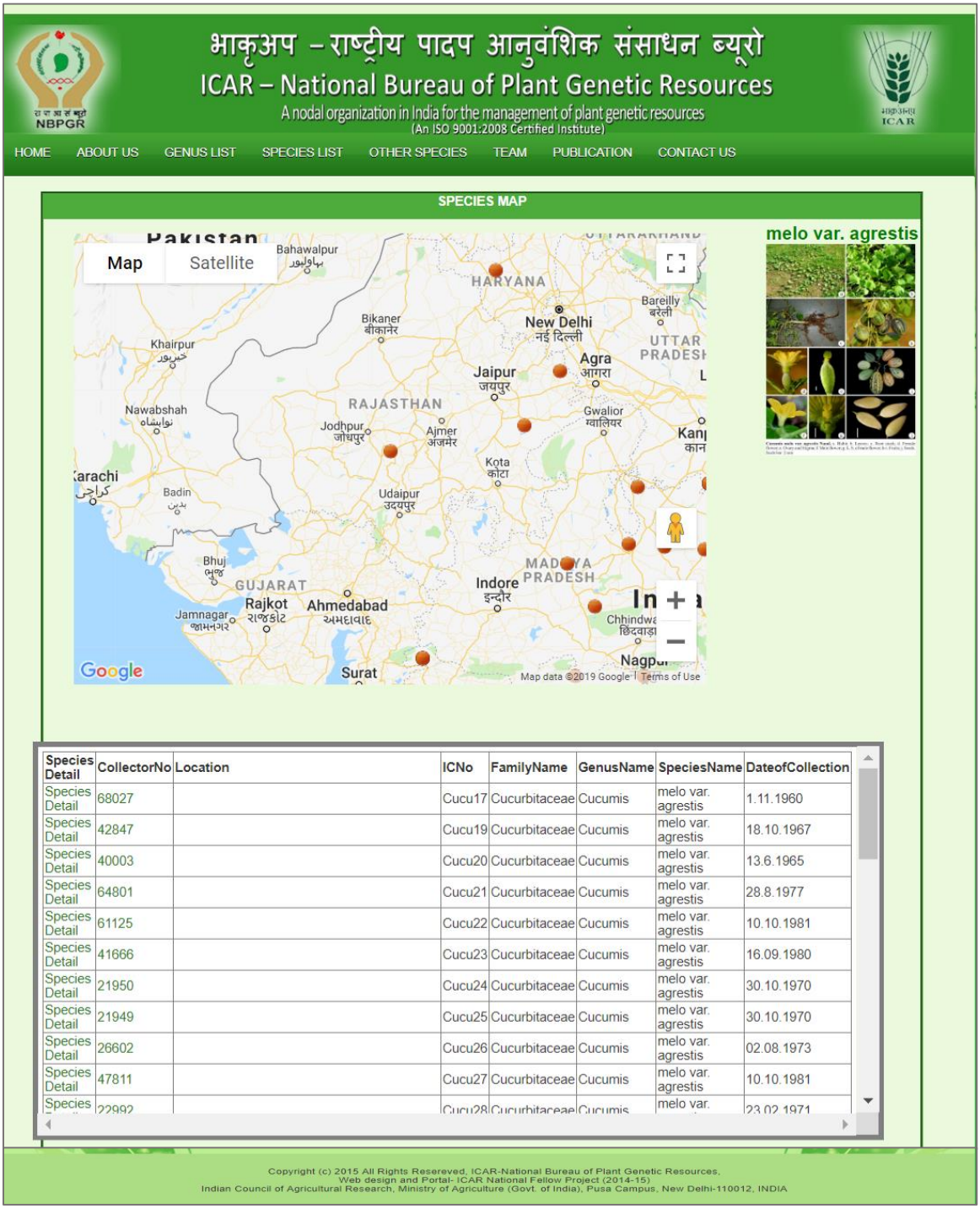
National genebanks face an uphill task of minimizing the gap between the germplasm collected by explorers and the actual numbers that ultimately find a place in the long term conservation set up. The challenge becomes stiffer in case of crop wild relatives (CWRs). Genebanks follow standards of conservations especially the minimum number of seeds whereas it is either difficult to establish the CWRs ex situ or takes multiple seasons to regenerate minimum number of seeds. Over a period of time the gap between collected and conserved yawns wide.

A web app was developed linking multiple databases to allow explorers to identify locations (passport data with geocodes) that require revisit and recollection for every crop.



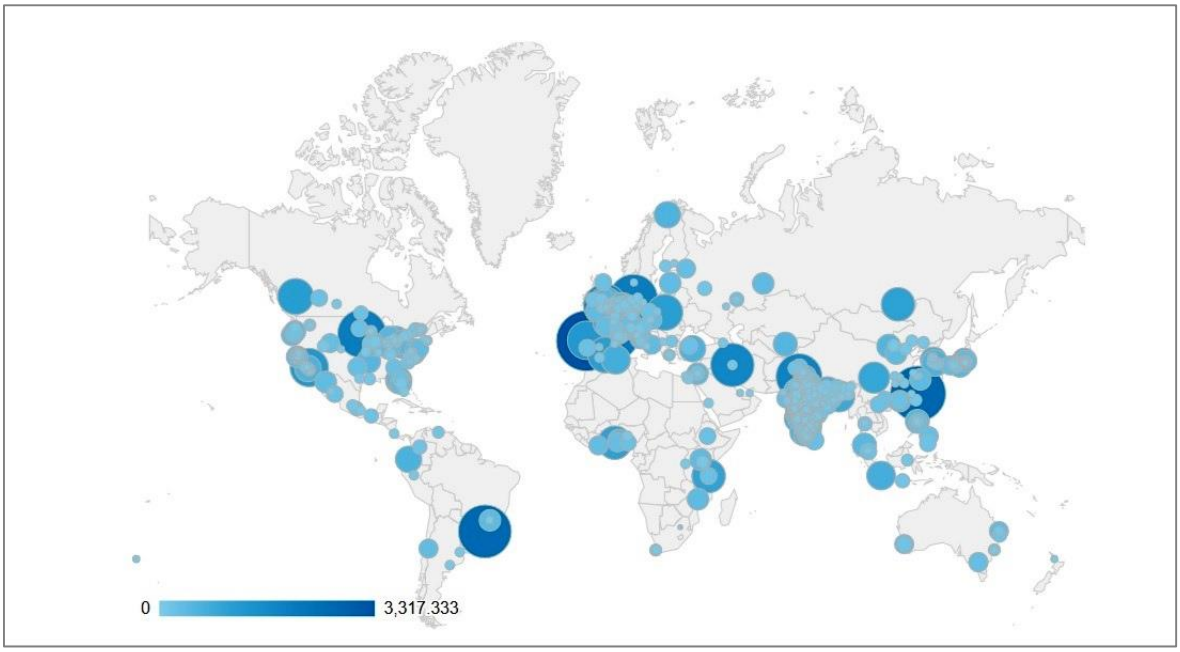
CWR Portal (http://pgrinformatics.nbpgr.ernet.in/cwr/)

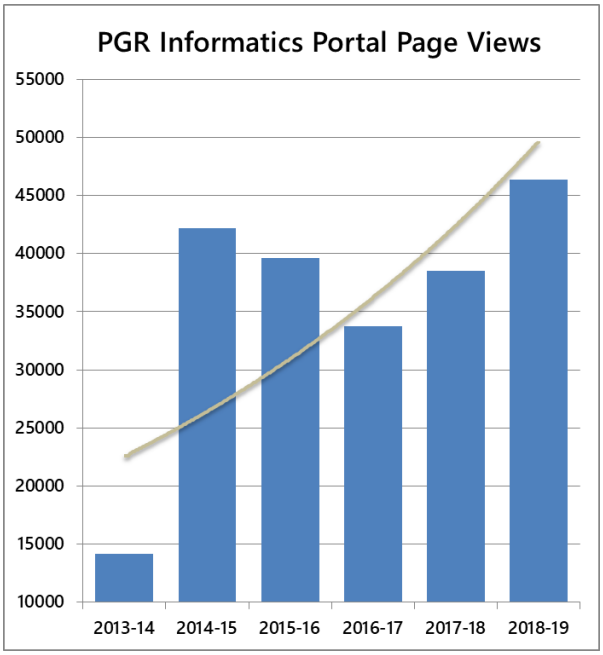
NAIP funded project "Biosystematics of the genera *Vigna*, *Cucumis* and *Abelmoschus*" operated by Dr. KV Bhat at NBPGR generated an exhaustive body of information in an effort to resolve species identities and reconstruct phylogeny among green gram, black gram, rice bean, adzuki bean and moth bean from the genus Vigna; cucumber, kakri, melons from the genus Cucumis and okra from genus Abelmoschus along with corresponding CWRs. However, it was not available for every user. Specialized databases were developed for taxonomy (including high resolution images to aid in identification) and passport data. An interactive web app was developed to provide taxonomic and distributional data to the researchers.



5. Impact of the Project

The ICAR-NF project on PGR Informatics set out to develop novel algorithms and to improvise those that were used in different fields with an aim to create modules that would enhance the efficiency of PGR management activities. Their applications were gradually incorporated into the PGR Portal that was in place since Dec 2012.





Usage statistics of PGR Informatics Portal (left) and locations from where accessed (map above).

Source: Google Analytics (1 Oct 2014 to 07 Aug 2019).

Note the impetus provided by the ICAR National Fellow Project adding value to PGR Portal with multiple web-apps and mobile-apps.

The novelty and utility of the web apps have seen them accessed by researchers across the globe. Impetus provided by the **web-apps and mobile-apps developed in the NF project to the use of PGR Portal by researchers in many countries demonstrates the immense impact of the ICAR-NF project.**

6. Capacity building

Teaching and Research Guide

I hold a dual faculty position at IARI, New Delhi:

- Faculty of Plant Genetic Resources, NBPGR
- Faculty of Bioinformatics, IASRI

I taught/continue to teach the following courses as Course Leader:

1. BI 504 Evolutionary Biology (2+1)
2. BI 624 Genome Wide Association Study (2+1)
3. PGR 507 Information Management in Plant Genetic Resources (2+1)
4. PGR 500 Biodiversity and Plant Genetic Resources (2+0)
5. PGS 503 Intellectual Property and its Management in Agriculture (1+0)

Research Students

Plant Genetic Resources

1.Mr. Shailendra Solanki, 10648, *Analysis of genetic variation in Artocarpus hirsutus (Wild Jack) collections from Western Ghats*

2.Ms. Shephalika Amrapali, 10854, *Olfactory, biochemical and molecular profiling of rose germplasm for fragrance*

3.Mr. Puneeth, 11304, *Development of Informatics System to Document on Farm Conservation: A Case Study*

Bioinformatics

4.Ms. Soumya Sharma, 10778, *Development of database of genes and gene families responsible for nutritional content in field crops*

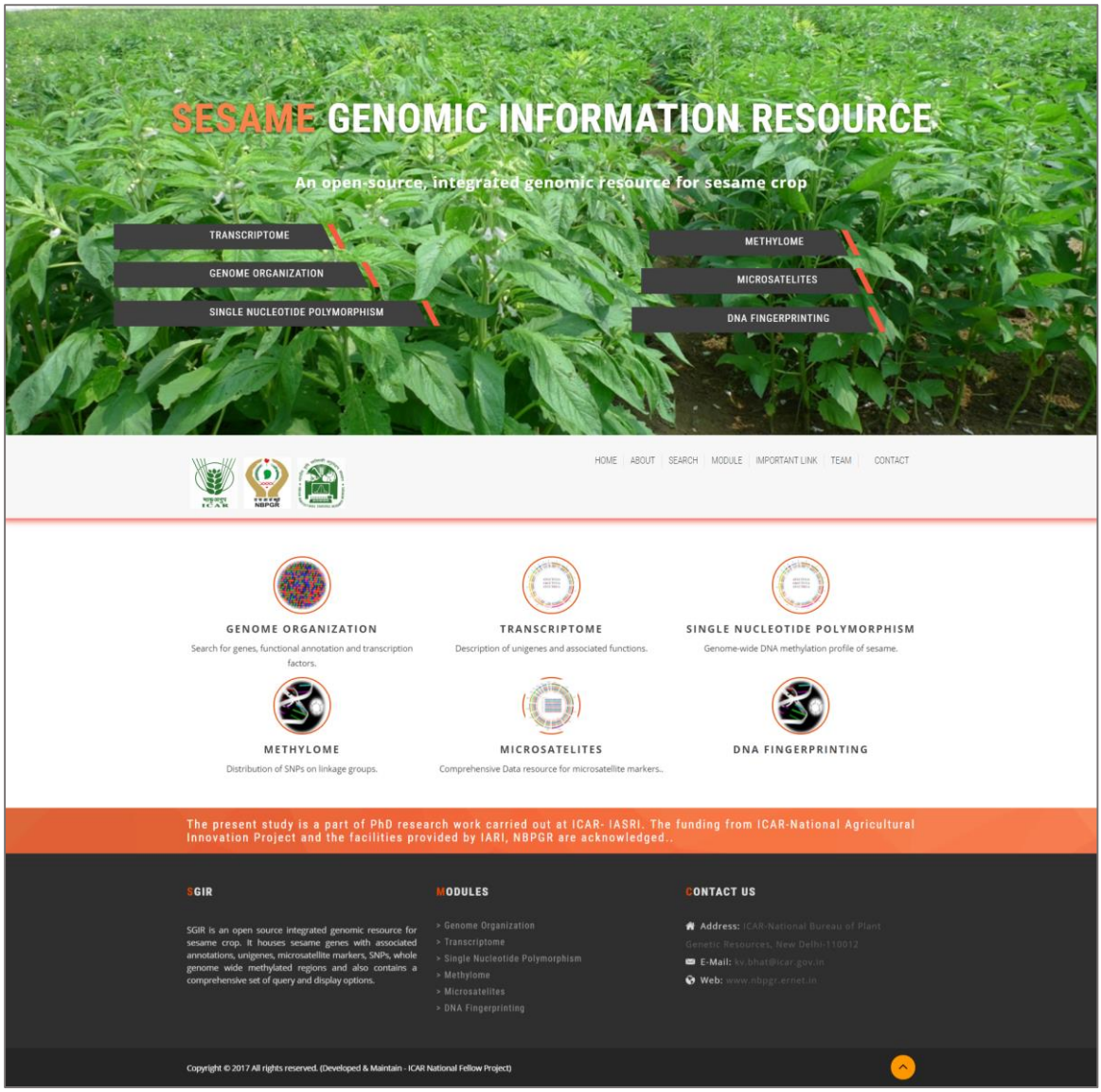
5.Ms. Sneha Murmu, 11006, *Computational approaches to understand host-pathogen interactions between wheat and its blast fungus*

6.Ms. Shweta Kumari, 11007, *Comparative Analysis of Domestication Related Genes in Minor Pulses*

7.Mr. Dipro Sinha, 11227, *Deep learning methods to link germplasm and genomic data*

Hands-on training:

One of the Bioinformatics Ph.D. students Dr. Supriya P. worked on "Development of Bioinformatics Framework for Enhanced Utilization of Genomic Resources in Sesame (*Sesamum indicum* L.)". She was trained as well as provided hands on assistance under the National Fellowship facility to develop **Sesame Genomic Information Resource** particularly the genome browser in *jbrowse*.

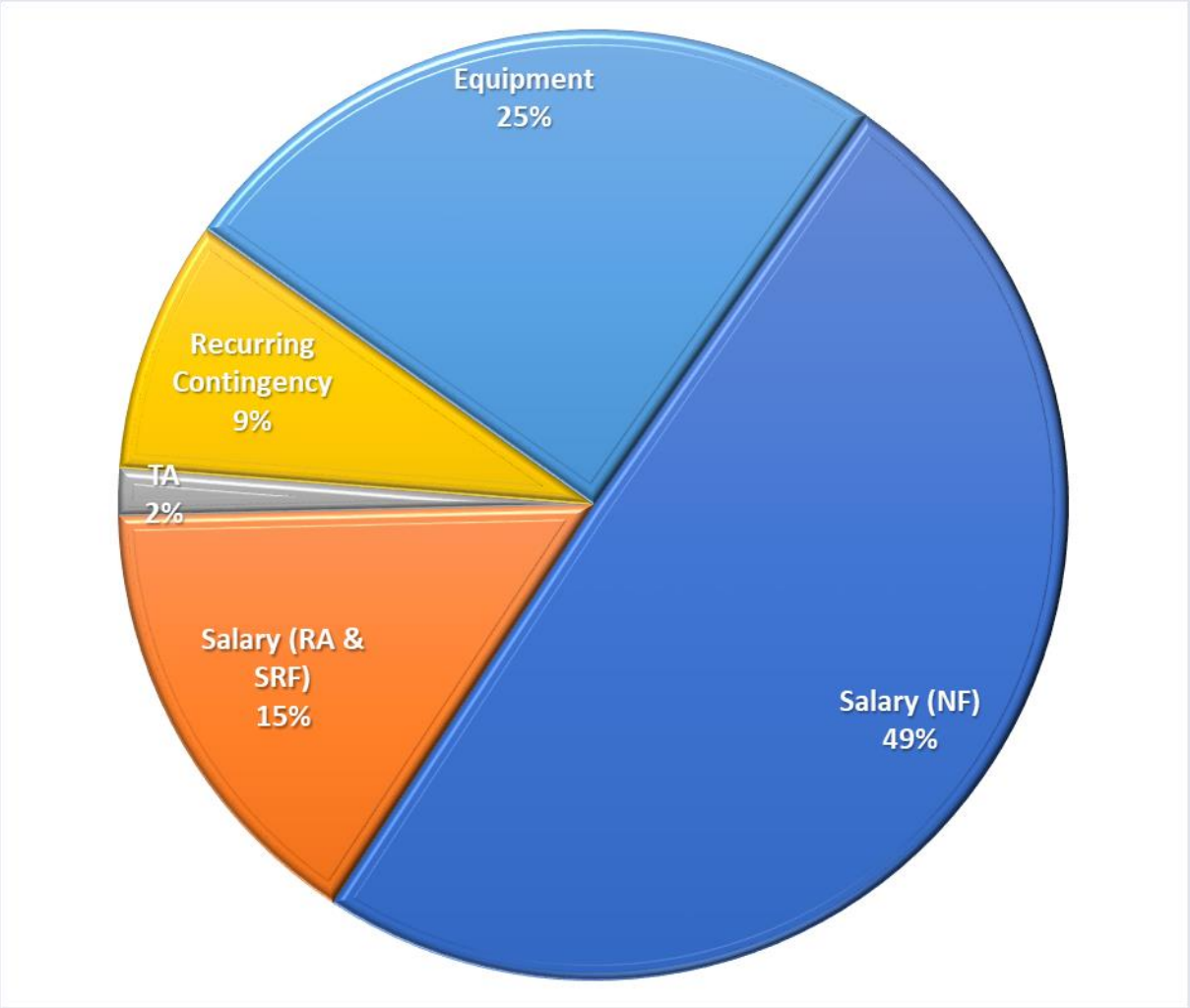


Organizing workshop/brainstorming meeting

1. A brainstorming meeting on "Strategies for Implementation of Delhi Declaration on Agrobiodiversity Management in India" was co-organized by NBPGR at New Delhi on August 28, 2017. The objective was to chalk out a plan for effective implementation of the 12-point Delhi Declaration on Agrobiodiversity Management. I contributed (as Chair, Technical Program, Proceedings and Publication Committee) in conceptualization, preparation of background notes and compiling and printing of proceedings that included action plan. One of the issues was PGR Informatics.

2. A "Regional Workshop on the Preparation of the National Reports on the Implementation of the International Treaty on PGRFA" was organized by the Secretariat of the International Treaty in collaboration with the Ministry of Agriculture and Farmers Welfare, Government of India, the FAO India and NBPGR (NASC, Pusa, New Delhi, from 11-13 December 2018). 25 Participants from thirteen countries (Afghanistan, Bangladesh, Bhutan, Fiji, Indonesia, India, Italy, Malaysia, Philippines, Mongolia, Philippines, Nepal, Thailand and Papua New Guinea) took part in the Workshop. I contributed as a local organizer as well as a technical expert in PGR Policy particularly PGR Informatics.

7. Budget



Budgetary allocation across heads (2015-2018). Figures based on releases up to 28-11-2019.